Voucher Specimen Project - Final

12 December 2003

Allan O'Connell, Principal Investigator
Andrew Gilbert, Research Technician
Patuxent Wildlife Research Center, USGS
11510 American Holly Drive, Laurel, MD 20708

GOALS

The project goal was to locate specimens in natural history collections that were collected within national park boundaries in 9 parks in the Northeast Temperate Network: Acadia National Park (ACAD), Marsh-Billings-Rockefeller National Historical Park (MABI), Minute Man National Historical Park (MIMA), Morristown National Historical Park (MORR), Roosevelt-Vanderbilt National Historic Site (ROVA), Saint-Gaudens National Historic Site (SAGA), Saugus Iron Works National Historic Site (SAIR), Saratoga National Historical Park (SARA), and Weir Farm National Historic Site (WEFA) and 5 parks in the Coastal and Barrier Island Network: Assateague Island National Seashore (ASIS), Cape Cod National Seashore (CACO), Fire Island National Seashore (FIIS), Gateway National Recreation Area (GATE), and Sagamore Hill National Historic Site (SAHI) (Table 1).

METHODS

Data requests

We obtained information about vertebrate (except fish) and vascular plant natural history collections having specimens from the eastern United States. We first searched two web-accessible databases of natural history collections: the Index Herbariorum (IH) (http://www.nybg.org/bsci/ih/ih.html) and the Directory of Research Systematics Collections (DRSC) (http://www.nbii.gov/datainfo/syscollect/drsc/). We also sent out requests for information about collections to several e-mail list-serves (TWS-L, NHCOLL-L, ORNITH-L) and obtained a list of museum contacts from John Karrish (NPS, Philadelphia) from a similar project. Additional collection information was found by searching the websites of regional biology departments.

Information about natural history collections was recorded in a Microsoft Access 2000 database. Collections were separated by taxa (e.g. Cornell University Museum of Vertebrates ornithology collection) where taxa-specific data were available. Information such as size of collection, percentage computerized, contact person and address, web address, and notes about the collections were recorded. We determined that much of the information provided in the two natural history collection databases were out of date; therefore, we checked contact information for all institutions through websites or by contacting institutions directly and updated information as necessary.

We mailed requests for data to 274 collection managers curating 299 natural history collections and 8 state natural heritage programs. We requested data for specimens originating

within 14 northeastern national parks in the Northeast Temperate Network (9) and Coastal and Barrier Island Network (5) (Table 1). To reduce search time and increase the number of responses from institutions, we broadened search criteria to county-wide locality requests. This approach also had the benefit of including locations that were miss-spelled or used historic names, which would not have otherwise been selected. We sent institutions a list of parks and localities by state and county(s). We requested that the following data fields be provided: park name, taxonomic name, common name, catalog number, accession number, condition of specimen, collector's name, date of collection, locality information, latitude-longitude, and comments. We e-mailed follow up requests for data to 177 collection managers who did not respond within 6 weeks of the initial request for data. Responses were logged into the collection database as they were received.

Searches of specimen collections

We searched 22 collections, 12 via the Internet and 10 manually. Those collections that were searched manually were selected from the pool of collections that were not (or only partially) computerized and from which we had not received data. We randomly selected five vertebrate and five plant collections from each of three size categories: small (≤30,000), medium (≤195,000), and large (>195,000) and subsequently chose at least one specimen collection from each of these categories. Size categories were determined by dividing the size distribution of collections in thirds. We searched herbaria systematically beginning with the first collection cabinet (or last in one case) and continuing through as much of the collection as possible during the time available. Where possible, we reduced the number of herbaria folders by searching only area specific folders (e.g., New England folders at Harvard University). Herbaria sheets were scanned for pertinent localities and data recorded directly into MS Excel spreadsheets. We searched the vertebrate collections at Chicago Academy of Sciences, Northeastern University, and London Museum of Natural History bird collection by scanning specimen tags, but searched the Harvard ornithological and London Museum of Natural History mammal collections by searching accession catalogs. All collections were arranged primarily taxonomically and secondarily geographically, allowing us to limit searching to appropriate geographic regions. We searched catalogs to reduce handling of specimens and increase search efficiency. However, taxonomy within catalogs were normally not updated and there was no assurance that specimens were still present in the collection. Most transfers to other institutions were noted in catalogs, but in some cases disposal of specimens was not recorded. We were unable to verify the presence of relevant specimens found in catalogs or corroborate the taxonomic identification in any collection due to time constraints.

Data manipulation

All data files were checked for errors and converted to Microsoft Excel 2000 format. Data were standardized to include the following data fields: institution, taxa, catalog number, original genus, original species, original subspecies or variety, family, updated genus, updated species, updated subspecies or variety, common name, state, county, specific locality, park, proximity to park (1 = within park boundaries, 2 = may be within park boundaries, 3 = in county, 4 = in state), collector, date collected, year collected, sex, age, parts/preparation, remarks, type status, latitude, longitude, and elevation. We accepted the taxonomic identification of all specimens without verification; however, we did update taxonomy to conform to current Integrated Taxonomic Information System (IT IS 2002) standards. Proximity to the nearest

national park was determined for each specimen based on the categories noted above using maps (DeLorme 1998, 2000), Mapquest on-line mapping software (Mapquest.com, Inc. 2002), and Topozone.com (Maps a la Carte, Inc. 2000). Category 1 records must have been determined to have been collected within current park boundaries, even if specimens were collected prior to park establishment. Specimens were assigned category 2 records if locality information indicated that records were collected within a town where the park resides in whole or in part, but for which we did not have enough specific locality information to classify as being inside or outside park boundaries. Category 2 specimens include all records with general township locality information, but little or no more specific information. For example, a record from Concord, Massachusetts was assigned category 2 for MIMA, but another record from Concord, Massachusetts from the Assabet River was definitely not collected within park boundaries and was assigned category 3. Records were assigned category 3 when we were able to determine that records were from within the county the park resides in, but definitely not within park boundaries. Category 4 records include all records from any other county within the state and were assigned to the closest (or only) park in that state. Some records were left unassigned because of insufficient locality data.

Statistical analyses

We determined if the number of category 1 and 2 records varied with the size or age of the park using linear regression analysis. Data were log transformed and the same analyses performed to determine if the relationships were logarithmic.

RESULTS

As of 31 March 2002, we received responses to requests for information from 212 collections and 5 natural heritage programs (71% of all queried). We assembled data from 78 collections kept at 52 institutions and the Maine Natural Areas program, 57 directly from collection managers, 11 from web-enabled searches, and 10 from manual searches (Table 2). We assembled 31,110 specimen records (Table 3) of which 4,745 (15%) are from within park boundaries (category 1) and an additional 4,552 (15%) may be from within park boundaries (category 2), but for which we do not have enough information to determine their exact location. The majority of the specimens, 20,224 specimens (65%), were from within the county (category 3) and the remaining 1,312 (4%) from within the state (category 4). We left 277 unassigned (<1%) because 1) we were unable to identify current locality based on a historic place name, 2) there were discrepancies in the locality data, or 3) they could not be assigned to any one park. Specimens are well distributed between 1890 to 2000 with notable gaps in collecting occurring in 1910's and 1950's, roughly corresponding with war-time activities (Figure 1). The number of category 1 and 2 records found for each park logarithmically increased with the log size of the park ($n = 14$, $p < 0.0001$, $r^2 = 0.73$) (Figure 2). The number of records did not increase linearly ($n = 14$, $p > 0.05$) (Figure 3) or logarithmically ($n = 14$, $p > 0.2$) (Figure 4) with the age of the park.

More than one-third of all records were from Acadia National Park, the largest and oldest park in this study. While Acadia National Park had the most cat. 1 and 2 specimens (4,615), CACO (2,180) had specimens in a greater number of institutions (29) (Table 4). Concentrations of category 1 and 2 specimens could be found in any institution, but most specimens were found

in the largest museums such as Smithsonian Institution, American Museum of Natural History, Cornell University, and Harvard University (Table 4).

We gathered the most specimens records for plants (13,048) followed by birds (10,056), mammals (5,276), and amphibians and reptiles (2,730). Within the four taxa, specimens were found in 260 families, 909 genera, and 2,055 species/species hybrids (Table 5). Plant specimens were the most diverse taxa with the greatest number of species/species hybrids within the most families and genera (Table 5). Species were further divided among innumerable subspecies and varieties, but were not tallied.

DISCUSSION

We were able to assemble a large and diverse group of specimens for 14 parks with minimal staff and limited resources. These data are an important record of biodiversity, providing an excellent baseline dataset from which to evaluate historic biodiversity in our nation's parks. Despite the wealth of data, however, there are limitations to this data. First, the data was not error-checked. Specimens were assumed to be correctly identified and all data recorded accurately. However, it is likely that some specimens were not properly identified to species and or have not been revisited following changes in taxonomy. Nevertheless, the vast majority of identifications are likely to be accurate. Secondly, some data such as locality and date can often be very general or missing altogether. Older specimens often were missing dates or had locality information specific only to a town or county. However, keeping these limitations in mind, the data is extremely valuable as a tool for exploring changes in biodiversity, particularly given the excellent distribution of records over the prior century.

Access to computer records of natural history specimens was critical to this project. We received most records from those collections that were computerized. Computer records allowed collection managers to complete searches in minutes, rather than days or weeks, that would be necessary to search some collections by hand. Whereas, manually searching specimen records added only 2517 (8%) of the records collected and took 21 days to complete. Fortunately, most institutions have some form of computer record system or will be establishing one. The continued computerization of specimen data will bring with it additional relevant specimens. In addition, many institutions are developing web-accessible interfaces for searching their computer catalogs. This allows end-users easier access to data and reduces time spent by collection managers filling requests for data. These sites often allow the downloading of data to spreadsheet formats and streamline data gathering. In the future, this work will become increasingly easy to accomplish and bring a wealth of often overlooked data to use.

While, we only assembled 8% of the total records manually, these searches were an important component of this project and allowed us to access data from specimens that would not have been accessible otherwise. Some institutions are years from computerization, not because of technological issues, but because of the time and resources necessary to enter the data. In order to retrieve data from these collections this work must be completed by hand. For example, Harvard University Herbaria has less than 1% of their collection computerized of 5,000,000 plant specimens; and within 4 days of searching we were able to locate greater than 800 specimens after searching an estimated 2% of the collection. Extrapolation would then suggest that there could be as many as 40,000 relevant specimens in that collection alone!

However, without manually searching this collection, these data would not be accessible for many more years.

Searching by hand is time consuming, but can yield valuable data not otherwise available if relying only on computer accessible records. Most natural history collections were unable to search collections themselves because of the lack of time and resources to fulfill such requests, which places the responsibility of searching understandably on the organization requesting the data. In order to be most efficient with time, we would recommend several things. First, be prepared. Knowledge of the localities searching for, including the historical names is very helpful in identifying relevant specimens. In addition, lists of potential species for that region can help narrow the search field, although care must be taken not to exclude rare, extinct, and vagrant species. Efficiency in searching is also important. We suggest searching specimen tags if the collection is sufficiently divided by locality. In most large collections, specimens were divided regionally into separate folders (plants) or trays (vertebrates). While the largest collections are daunting in size, they often were the easiest to search because they possessed enough specimens to be divided into smaller regions. Smaller collections tended to be divided into local specimens, the rest of North America, and foreign specimens, requiring searching of most all of the specimens. While searching specimen tags can be tedious, searching by specimens has the advantage of having updated taxonomy and the assurance that specimens are still in the collection. But, specimen tags are often very difficult to read, particularly for vertebrate specimens with small tags and old writing. Additionally, handling specimens degrades the specimens and may be irritating to the searcher because of the use of harsh chemicals used for their preservation. Searching by catalog is much faster, but provides less reliable data and is usually not updated taxonomically. If time permits, searching through catalogs first, then referencing against specimens in the collection may be a good compromise. Ultimately, every collection is managed differently, which will affect how you choose to search a collection. Being flexible in strategy is important for determining the best method for conducting searches.

REFERENCES

DeLorme. 1998. The Maine atlas and gazateer. Yarmouth, ME.

DeLorme. 2000. Topographic Map of Mt. Desert Island & Acadia National Park, Hancock County, Maine. Yarmouth, ME.

Integrated Taxonomic Information System (IT IS) on-line database, http://www.itis.usda.gov, retrieved January 5- January 15, 2001.

Mapquest.com, inc. 2002. http://www.mapquest.com, retrieved January 15-February 11, 2002.

Maps a la Carte, Inc. 2000. http://www.topozone.com, retrieved January 15-February 11, 2002.

Table 1. National parks searched for vertebrate and vascular plant voucher specimens.

National Park (Code)	State(s)	Size (Ac)	Year Est.
Northeast Temperate Network			
Acadia National Park (ACAD)	ME	46784	1916
Marsh-Billings-Rockefeller National Historical Park (MABI)	VT	555	1992
Minute Man National Historical Park (MIMA)	MA	967	1959
Morristown National Historical Park (MORR)	NJ	1685	1933
Roosevelt-Vanderbilt National Historic Site (ROVA)[1]	NY	683	1940
Saint-Gaudens National Historic Site (SAGA)	NH	150	1964
Saugus Iron Works National Historic Site (SAIR)	MA	9	1968
Saratoga National Historical Park (SARA)	NY	3406	1938
Weir Farm National Historic Site (WEFA)	CT	60	1990
Coastal and Barrier Island Network			
Assateague Island National Seashore (ASIS)	MD	39732	1965
Cape Cod National Seashore (CACO)	MA	43604	1961
Fire Island National Seashore (FIIS)	NY	19580	1981
Gateway National Recreation Area (GATE)	NY, NJ	26610	1997
Sagamore Hill National Historic Site (SAHI)	NY	83	1963

[1] ROVA was consolidated from Eleanor Roosevelt National Historic Site (ELRO, est. 1977, 181 ac), Home of Franklin D. Roosevelt National Historic Site (HOFR, est. 1945, 290 ac) and, Vanderbilt Mansion National Historic Site (VAMA, est. 1940, 212 ac).

Table 2. Search results for herbaria and vertebrate collections manually searched in this study.

Collection searched	Size (category)[1]	Search time	No. records	Approx. % searched[2]
Herbaria				
Mary Washington College	5,000 (small)	4 hours	1	100
University of South Carolina	85,000 (medium)	3 days	54	100
University of Minnesota	818,000 (large)	4 days	50	17
Harvard University	5,000,000 (large)	4 days	817	2
University of Maryland	67,000 (medium)	2 days	288	33
Vertebrate				
Chicago Acad. Sciences – herp.	20,000 (small)	2 hours	16	25
Northeastern Univ. – vert.	42,000 (medium)	4 hours	148	5
Harvard Univ. – ornith.	338,000 (large)	3 days	504	66
London Mus. Nat. Hist. - birds	2,500,000 (large)	3 days	461	90
London Mus. Nat. Hist. - mammals	359,000 (large)	1 day	178	100

[1] Small ≤30,000, medium ≤195,000), and large >195,000.

[2] Estimated for herbaria based on the number of cabinets searched. Estimates for Chicago depend on the number of rows of specimens searched, for Northeastern are based on the number of cabinet drawers and catalog numbers, and for Harvard are based on catalog numbers. We received help searching from two other people at Chicago and one other at Northeastern.

Table 3. The number of specimen records received in each proximity category for all parks.

	Number of specimen records[1]				
Park code	Category 1	Category 2	Category 3	Category 4	Total (%)[2]
ACAD	3,392	1,223	7,739	149	12,580 (22.5)
MABI	1	199	273	20	493 (1.6)
MIMA	72	408	1,797	78	2,385 (7.7)
MORR	0	119	905	46	1,072 (3.4)
ROVA	237	4	251	485	977 (3.1)
SAGA	0	10	102	19	131 (0.4)
SAIR	0	17	722	0	739 (2.4)
SARA	180	6	115	423	724 (2.3)
WEFA	12	15	983	8	1,018 (3.3)
ASIS	471	1	197	3	672 (2.2)
CACO	186	1,994	1,806	6	4,010 (12.9)
FIIS	109	276	4,026	0	4,471 (14.4)
GATE	30	277	1,107	75	1,492 (4.8)
SAHI	55	3	201	0	259 (0.8)
Total (%)	4,745 (15.3)	4,552 (14.6)	20,224 (65.0)	1,312 (4.2)	31,110

[1] Category 1 = within park boundaries, 2 = may be within park boundaries, 3 = in county, 4 = in state.
[2] Totals are reduced by 277 specimens (0.89%), because we were unable to identify current locality based on a historic place name, there were discrepancies in the locality data, or they could not be assigned to any one park.

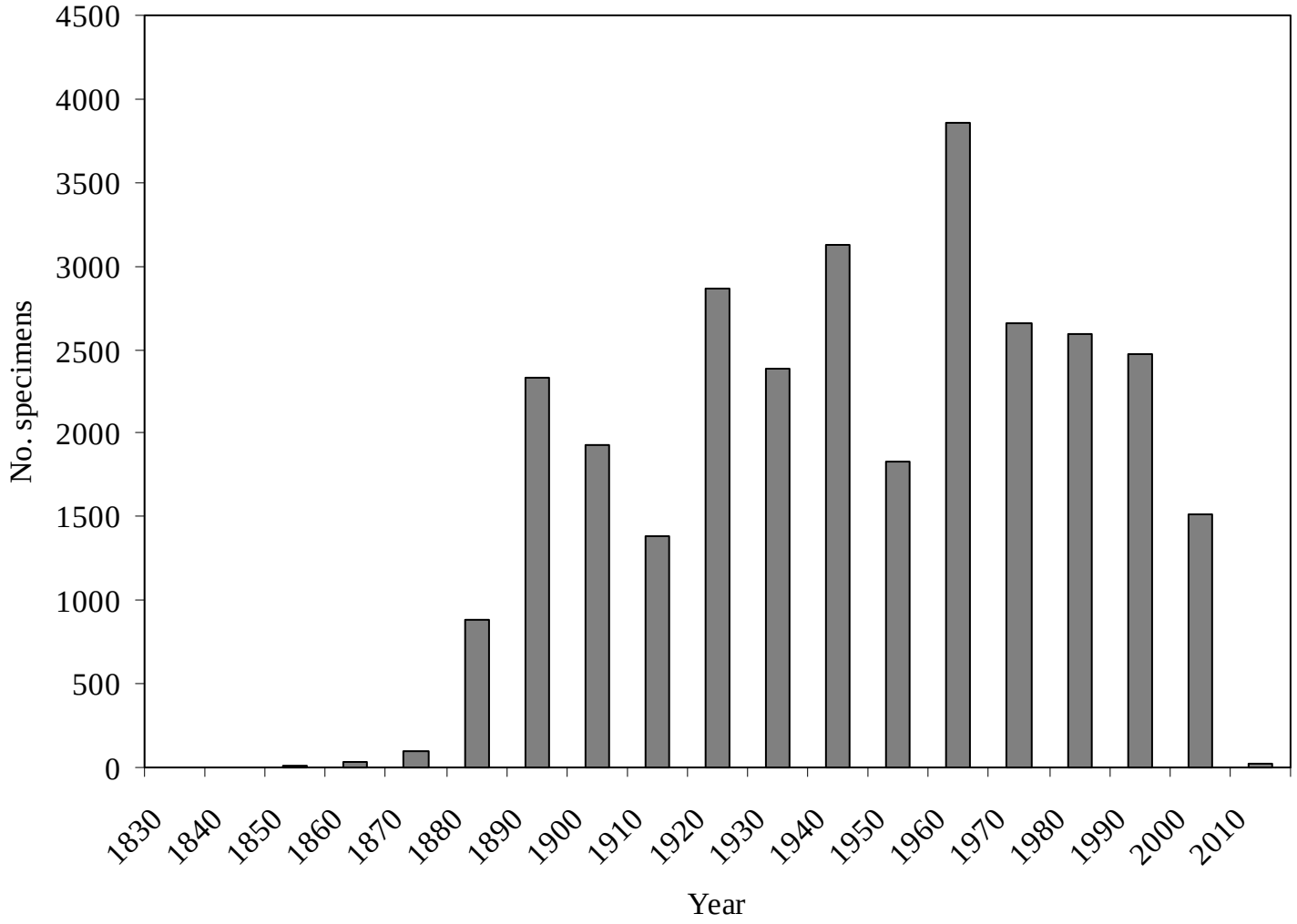

Figure 1. Distribution of year of collection for specimens for the study of vouchers specimens in the northeast. Notable absences in collecting occurred during the 1910's and 1950's roughly corresponding with periods of wartime activities.

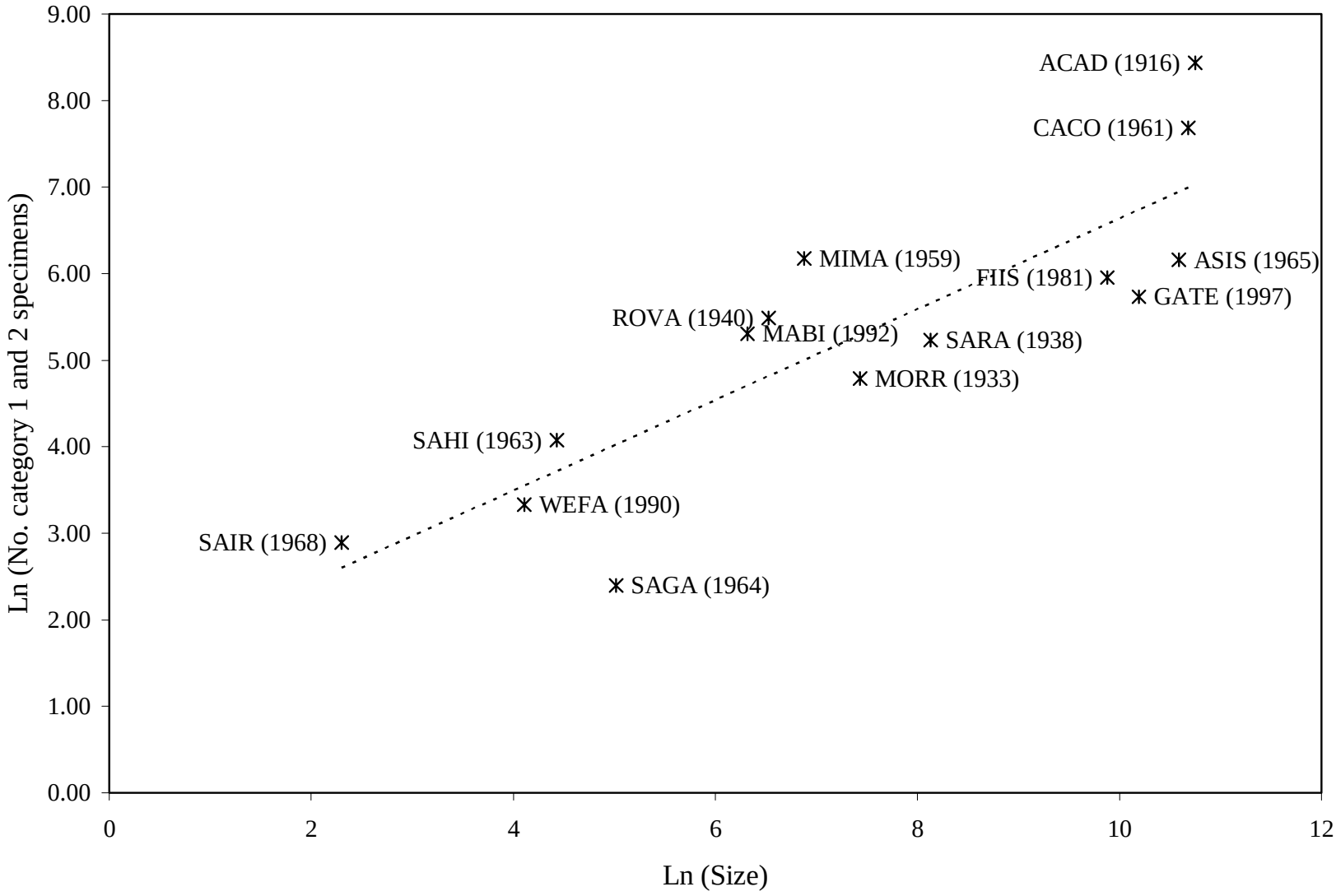

Figure 2. The natural log of the number of category 1 and 2 specimens vs. the natural log size of the park (in acres) in which they are associated. Parks are identified by park codes and includes the year established.

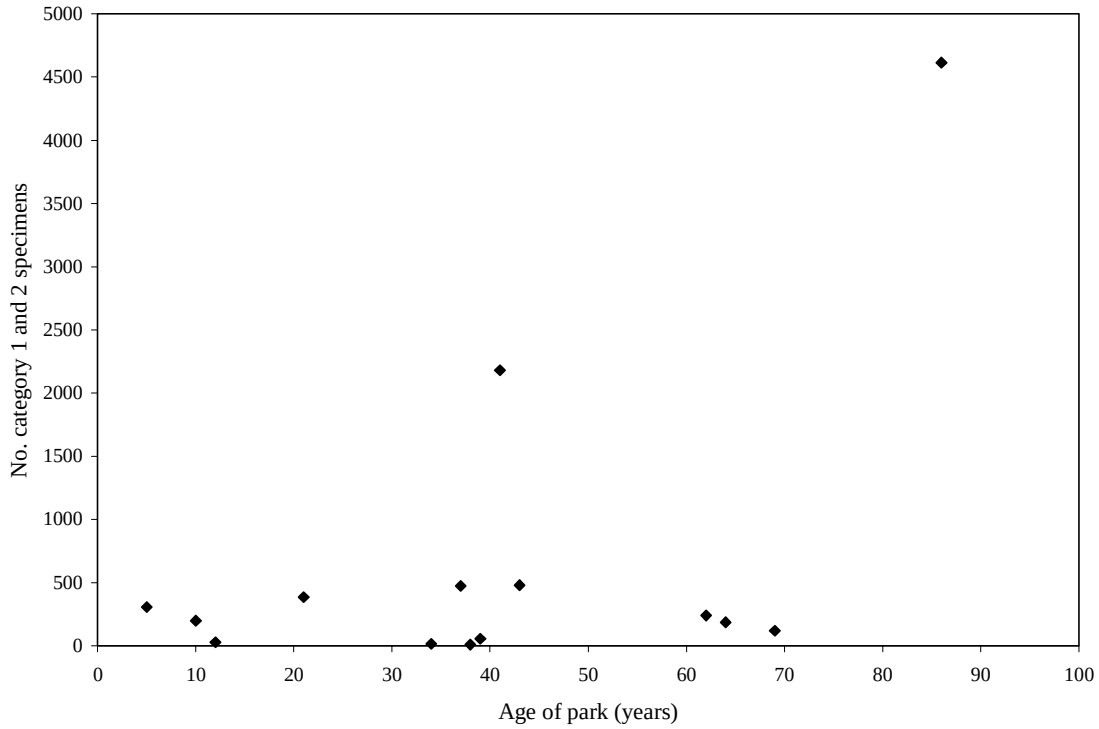

Figure 3. The number of category 1 and 2 specimens vs. the age of the park in which they are associated.

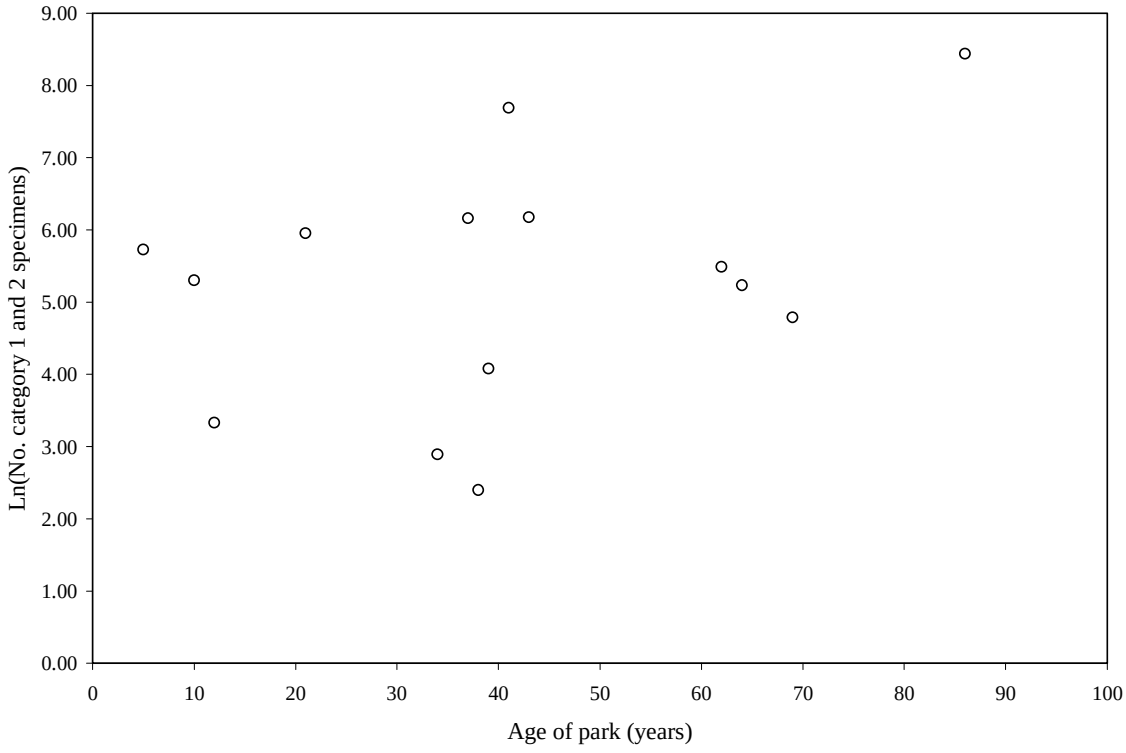

Figure 4. The natural log of the number of category 1 and 2 specimens vs. the age of the park in which they are associated.

Table 4. The number of institutions and the top two institutions with category 1 and 2 data for parks in this study.

Park code	No. institutions with cat. 1 and 2 data	Institution with greatest no. cat. 1 and 2 (no.)	Institution with 2nd greatest no. cat. 1 and 2 (no.)
ACAD	24	College of the Atlantic (3110)	University of Maine Herbaria (746)
MABI	8	Yale University, Peabody Museum of Natural History (78)	University of California, Berkeley, Museum of Vertebrate Zoology (63)
MIMA	18	Harvard University, Museum of Comparative Zoology (193)	Northeastern University Vertebrate Collection (66)
MORR	9	American Museum of Natural History (48)	Cornell University Museum of Vertebrates (26)
ROVA	5	American Museum of Natural History (160)	New York State Museum (70)
SAGA	1	American Museum of Natural History (10)	None
SAIR	5	Harvard University, Museum of Comparative Zoology (9)	Harvard University Herbaria (5)
SARA	5	New York State Museum (157)	Museum of Southwestern Biology (21)
WEFA	2	American Museum of Natural History (13)	University of Connecticut, EEB (9)
ASIS	8	University of Maryland, Norton-Brown Herbarium (288)	Smithsonian Institution, National Museum of Natural History (167)
CACO	29	Natural History Museum, London (597)	Harvard University, Museum of Comparative Zoology (401)
FIIS	14	Cornell University Museum of Vertebrates (204)	American Museum of Natural History (78)
GATE	20	Smithsonian Institution, National Museum of Natural History (86)	American Museum of Natural History (55)
SAHI	4	University of South Carolina, A. C. Moore Herbarium (51)	Natural History Museum, London (4)

Table 5. Taxonomic diversity and total number of category 1 to 4 specimens.

Taxa	No. Specimens	Family	Genera	Spp./hybrids
Birds	10,056	59	196	353
Herps	2,730	18	46	75
Mammals	5,276	32	75	111
Plants	13,048	151	592	1,516
Total	31,110	260	909	2,055